MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 3, 2017/2018

BMD3024 – MULTIVARIATE DATA ANALYSIS

(All sections / Groups)

31 MAY 2018
9.00 a.m. - 11.00 a.m.
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of **EIGHT (8)** printed pages with **FOUR (4)** structured questions.
2. Marks are shown at the end of each question.
3. Write all your answers in the answer booklet provided.
4. Answer **ALL** questions.

Question One (25 marks)

(a) Classify the level of measurement of the following variables. Explain your classification.

(i) The number of downloads of different bands' songs on iTunes.
(2 marks)

(ii) The names of the bands downloaded.
(2 marks)

(iii) The position of artistes and bands in the iTunes download chart.
(2 marks)

(iv) The money earned by the bands from fans' downloads from iTunes.
(2 marks)

(b) Summarize all the key assumptions as well as other assumptions of Multiple Discriminant Analysis.
(6 marks)

(c) Discuss the advantages of Multivariate Analysis of Variance (MANOVA), besides being able to analyse multiple dependent variables.
(6 marks)

(d) Describe the pros and cons of **TWO (2)** Hierarchical Clustering methods.
(5 marks)

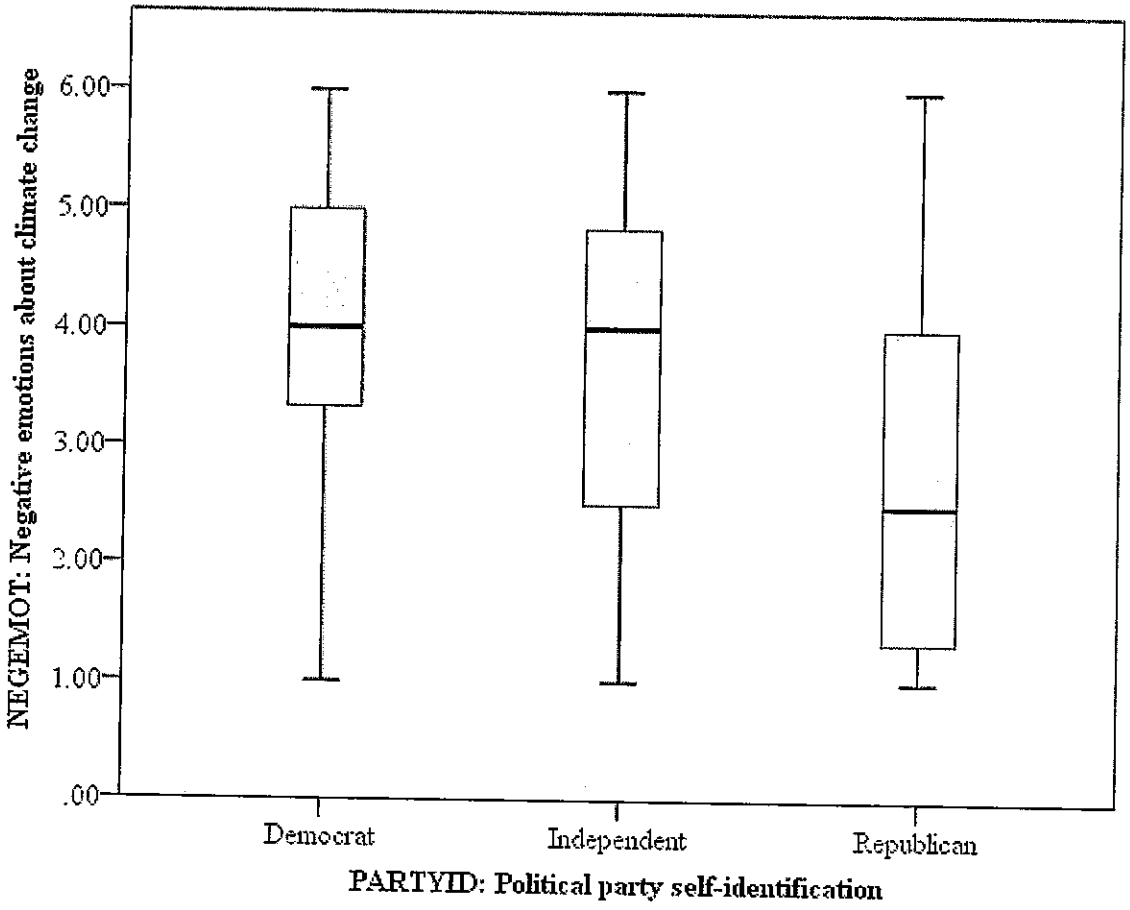
Continued...

Question Two (25 marks)

(a) Analyse the boxplot diagram in **Figure 1**.

(8 marks)

Figure 1: Boxplot Diagram for Negative Emotions About Climate Change According to Political Party Self-Identification



Continued...

(b) A study was carried out to explore the relationship between Number of Visits to Health Professionals (*timedrs*) and several potential predicting factors in 465 respondents. Variables measured were Physical Health Symptoms (*phyheal*) (high score = poorer physical health), Mental Health Symptoms (*menheal*) (high score = poorer mental health), and Stressful Life Events (*stress*) (high score = more stressful events). The researcher conducted a Multiple Regression and obtained the following output.

Table 1: Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.468[a]	.219	.214	9.708	2.007

a. Predictors: (Constant), *stress*, *phyheal*, *menheal*

b. Dependent Variable: *timedrs*

Table 2: ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12168.315	3	4056.105	43.034	.000[b]
	Residual	43451.134	461	94.254		
	Total	55619.449	464			

a. Dependent Variable: *timedrs*

b. Predictors: (Constant), *stress*, *phyheal*, *menheal*

Table 3: Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	-3.705	1.124		-3.296	.001		
phyheal	1.787	.221	.390	8.083	.000	.729	1.372
menheal	-.010	.129	-.004	-.075	.940	.694	1.441
stress	.014	.004	.169	3.769	.000	.844	1.184

a. Dependent Variable: timedrs

Continued...

(i) Interpret **Table 1**. Why is the adjusted R^2 lower than the unadjusted R^2?

(6 marks)

(ii) Examine **Table 2**. Is the *F*-statistic significant at 5 percent level? What does this indicate?

(4 marks)

(iii) According to **Table 3**, which are the variables that are significant in the prediction at 5 percent level?

(2 marks)

(iv) According to **Table 3**, which are the variables with worse multicollinearity scores than the others?

(5 marks)

Question Three (25 marks)

(a) Dr. Stephen Strange decided to compare his patients against a normal sample. He observed 10 of his patients as they went through a normal day. He also observed 10 lecturers at his university. The participants' category is named *Group*. He measured all participants using two dependent variables: how many chicken impersonations they did (*Quantity*), and, how good their impersonations were (as scored out of 10 by an independent farmyard noise expert) (*Quality*). They conducted a MANOVA analysis using SPSS and obtained the following output.

Table 4: Box's Test of Equality of Covariance Matrices[a]

Box's M	20.926
F	6.135
df1	3
df2	58320.000
Sig.	.000

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + group

Continued...

Table 5:Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[c]
Intercept	Pillai's Trace	.919	96.201[b]	2.000	17.000	.000	.919	192.402	1.000
	Wilks' Lambda	.081	96.201[b]	2.000	17.000	.000	.919	192.402	1.000
	Hotelling's Trace	11.318	96.201[b]	2.000	17.000	.000	.919	192.402	1.000
	Roy's Largest Root	11.318	96.201[b]	2.000	17.000	.000	.919	192.402	1.000
Group	Pillai's Trace	.333	4.250[b]	2.000	17.000	.032	.333	8.500	.662
	Wilks' Lambda	.667	4.250[b]	2.000	17.000	.032	.333	8.500	.662
	Hotelling's Trace	.500	4.250[b]	2.000	17.000	.032	.333	8.500	.662
	Roy's Largest Root	.500	4.250[b]	2.000	17.000	.032	.333	8.500	.662

a. Design: Intercept + group

b. Exact statistic

c. Computed using alpha = .05

(i) Refer to **Table 4** above. Using 5 percent level of significance, was the assumption of homogeneity of covariance violated?

(4 marks)

(ii) Based on the output in **Table 5**, is the variable *Group* a significant predictor for the dependent variate, using Hotelling's Trace statistic? Use 5 percent level of significance.

(4 marks)

Continued...

Table 6: Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parame-ter	Obser-ved Power[c]
Correct-ed Model	Quality	4.05[a]	1	4.050	.806	.381	.043	.806	.136
	Quantity	9.80[b]	1	9.800	.530	.476	.029	.530	.106
Inter-cept	Quality	1022.45	1	1022.45	203.360	.000	.919	203.360	1.000
	Quantity	2599.20	1	2599.200	140.497	.000	.886	140.497	1.000
Group	Quality	4.05	1	4.050	.806	.381	.043	.806	.136
	Quantity	9.80	1	9.800	.530	.476	.029	.530	.106
Error	Quality	90.50	18	5.028					
	Quantity	333.00	18	18.500					
Total	Quality	1117.00	20						
	Quantity	2942.00	20						
Correc-ted Total	Quality	94.55	19						
	Quantity	342.80	19						

a. R Squared = .043 (Adjusted R Squared = -.010)

b. R Squared = .029 (Adjusted R Squared = -.025)

c. Computed using alpha = .05

(iii)Based on the output in **Table 6**, analyse and report the results of the univariate effects of *Group* on each of the dependent variables at 5% level of significance. (6 marks)

Continued...

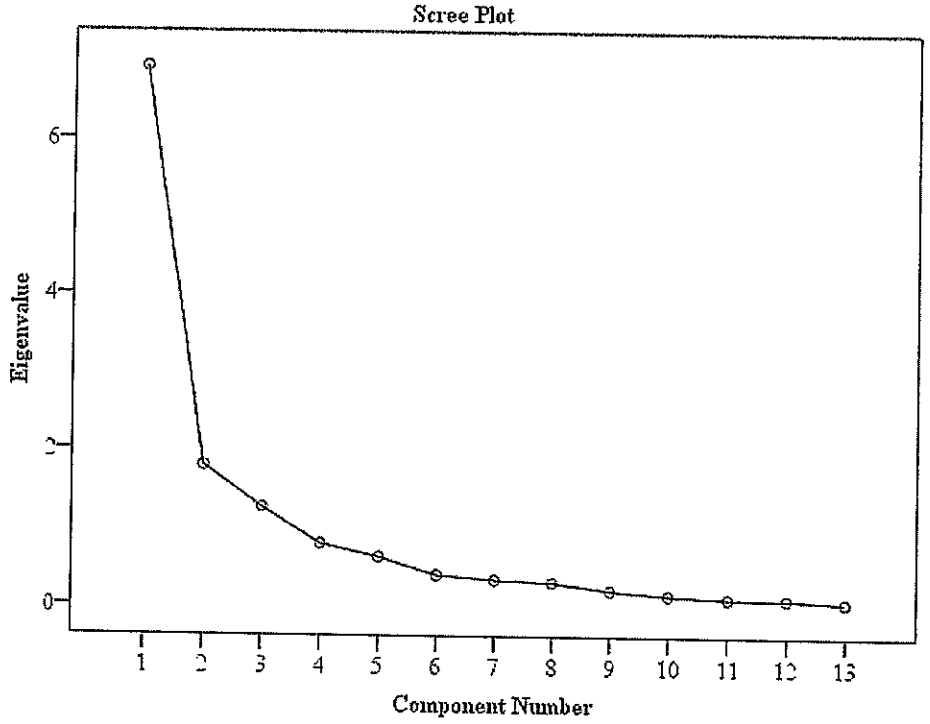
(b) **Table 7** and **Figure 2** are the SPSS output of an Exploratory Factor Analysis.

Table 7: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.926	53.277	53.277	6.926	53.277	53.277	3.611	27.777	27.777
2	1.789	13.758	67.035	1.789	13.758	67.035	3.534	27.183	54.959
3	1.253	9.638	76.673	1.253	9.638	76.673	2.823	21.714	76.673
4	.786	6.044	82.717						
5	.623	4.795	87.511						
6	...	.	...						
7									

Extraction Method: Principal Component Analysis.

Figure 2



Continued...

(i) Based on **Table 7**, use the latent root criterion of retaining factors to determine and explain how many factors will be retained.

(4 marks)

(ii) Based on **Figure 2** use the Scree Test criterion of retaining factors to determine and explain how many factors will be retained.

(4 marks)

(iii)Based on your results in (i) and (ii), explain how many factors will be retained.

(3 marks)

Question Four (25 marks)

(a) Develop one (1) path diagram that shows the following hypothesized structural relationships amongst four (4) constructs: *F1*, *F2*, *F3*, and *F4*; and two (2) mediating variables, *M1* and *M2*.

- *F1*, *F2*, and *F3* are independent factors.
- *F4* is a dependent factor.
- *F1* is made of five (5) variables.
- *F2* to *F4* are made of four (4) variables each.
- *M1* mediates the relationship between *F1* and *F4*.
- *M2* mediates the relationship between *F2* and *F4*.

(20 marks)

(b) Label all the following hypotheses and relationships between factors in the diagram:

H_1: *F1* is positively related to variable *M1*.
H_2: *F2* is positively related to variable *M2*.
H_3: *M1* is positively related to *F4*.
H_4: *M2* is negatively related to *F4*.
H_5: *F3* is positively related to *F4*

(5 marks)

End of Page